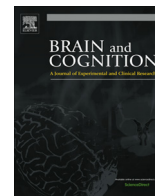




Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

The neural correlates of moral decision-making: A systematic review and meta-analysis of moral evaluations and response decision judgements



Beverley Garrigan^a, Anna L.R. Adlam^b, Peter E. Langdon^{c,d,*}

^a Department of Clinical Psychology, Norwich Medical School, University of East Anglia, United Kingdom

^b School of Psychology, College of Life and Environmental Sciences, University of Exeter, United Kingdom

^c Tizard Centre, University of Kent, United Kingdom

^d Hertfordshire Partnership University NHS Foundation Trust – Norfolk, United Kingdom

ARTICLE INFO

Article history:

Received 16 March 2016

Revised 14 July 2016

Accepted 25 July 2016

Keywords:

Moral evaluation (ME)

Moral response decision (MRD)

Activation likelihood estimate analysis (ALE)

Meta-analysis

Functional magnetic resonance imaging (fMRI)

ABSTRACT

The aims of this systematic review were to determine: (a) which brain areas are consistently more active when making (i) moral response decisions, defined as choosing a response to a moral dilemma, or deciding whether to accept a proposed solution, or (ii) moral evaluations, defined as judging the appropriateness of another's actions in a moral dilemma, rating moral statements as right or wrong, or identifying important moral issues; and (b) shared and significantly different activation patterns for these two types of moral judgements. A systematic search of the literature returned 28 experiments. Activation likelihood estimate analysis identified the brain areas commonly more active for moral response decisions and for moral evaluations. Conjunction analysis revealed shared activation for both types of moral judgement in the left middle temporal gyrus, cingulate gyrus, and medial frontal gyrus. Contrast analyses found no significant clusters of increased activation for the moral evaluations-moral response decisions contrast, but found that moral response decisions additionally activated the left and right middle temporal gyrus and the right precuneus. Making one's own moral decisions involves different brain areas compared to judging the moral actions of others, implying that these judgements may involve different processes.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past decade, functional magnetic resonance imaging (fMRI) has increasingly been used to measure the neural correlates of moral decision-making, adding to our understanding of the cognitive and affective processes involved. Nevertheless, there are issues with a lack of consistency amongst studies (Christensen & Gomila, 2012); a variety of different tasks have been used, and there are no agreed definitions, meaning that moral terms such as judgement, reasoning, sensitivity and moral cognition are all used differently across experiments. For the purpose of this study, we define moral judgements from a developmental psychology perspective; a moral judgement can refer to any judgement made within the moral domain, i.e. judgements relating to moral principles such as harm, justice, and fairness (Smetana, 2006; Turiel, 1983). Moral judgements can either be response decisions about what to do in a moral dilemma (self), or can be judgements of others, including judging individuals, groups, institutions or moral

principles. The distinction between different types of moral judgement has not been explicitly recognised amongst cognitive neuroscientists, with recent meta-analyses in this field grouping all task types together when analysing the neural correlates of moral decision-making. Task-type may influence the results, and whether moral judgements related to the self involve different processes, and different brain areas relative to moral judgements about others has not yet been considered in previous systematic reviews.

The moral tasks used in fMRI experiments which involve an active judgement (as opposed to passive judgements) can be grouped into two categories: (a) moral response decision tasks, where an individual is asked to make a decision (judgement) about what they would do in a hypothetical moral dilemma; and (b) moral evaluation tasks, where an individual is asked to judge the appropriateness or moral permissibility of another's actions, or asked to identify or judge a moral issue or violation. Moral response decisions require an individual to think about what they would do in a moral dilemma, whereas moral evaluations require judging the moral permissibility or appropriateness of the actions of others in a moral dilemma. Discrepancy has been found for answers to moral response questions ("Would you do X?") and moral evaluation questions ("Is it wrong to do X?") in moral

* Corresponding author at: Tizard Centre, University of Kent, Canterbury CT2 7LR, United Kingdom.

E-mail address: P.E.Langdon@kent.ac.uk (P.E. Langdon).

dilemmas (Tassy, Oullier, Mancini, & Wicker, 2013); “it seems that deciding *what to do* is not processed in the same way as deciding whether an action is *right or wrong*, and that in moral dilemmas it is the first that matters” (Christensen, Flexas, Calabrese, Gut, & Gomila, 2014, p. 5). A systematic review of moral decision-making which compares brain activation patterns for moral evaluations and moral response decisions can help to address whether these different questions are indeed processed in different ways in the brain. We hypothesised that making one’s own decisions about what to do in a moral dilemma and judging the moral actions of others will show increased activation of different brain areas, with response decisions showing greater activation in self-referential regions and evaluations showing greater activation in theory of mind (ToM) regions.

There appears to be no evidence for a uniquely “moral brain” (Young & Dungan, 2012), as brain areas that show increased activation during moral tasks are also involved in other functions. However, the brain region which appears to be of particular importance for morality, based on neuroimaging and lesion studies, is the ventromedial prefrontal cortex (vmPFC Blair, Marsh, Finger, Blair, & Luo, 2006; Fumagalli & Priori, 2012; Marazziti, Baroni, Landi, Ceresoli, & Dell’Osso, 2013; Raine & Yang, 2006). The vmPFC is thought to be involved in emotion regulation, and activation during moral decision-making tasks is seen as evidence of the involvement of emotion processes in making moral judgements (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). Reviews of moral neuroimaging evidence have also suggested that ToM is a key cognitive input to moral judgement because ToM brain regions show increased activation when making moral judgements (Blair et al., 2006; Young & Dungan, 2012). However, this conclusion may have been overstated because most neuroimaging experiments utilise moral evaluation tasks, where participants are asked to evaluate the actions of *others*. So while ToM is likely to be involved in judging the moral permissibility or appropriateness of others’ actions, it remains to be seen whether ToM brain regions are as active when making one’s own moral response decisions.

Two recent meta-analyses have been conducted on brain areas consistently showing increased activation in moral decision-making studies. Bzdok et al. (2012) performed an activation likelihood estimate (ALE) analysis of morality, empathy and ToM and found overlap in activation for ToM and morality. Experiments were only included in the ‘moral cognition’ domain, which they defined as a “reflection of the social appropriateness of people’s actions” (p. 789) if the task required participants to make judgements of other people’s actions. It is, therefore, not surprising that there was overlap with ToM brain activation, as moral evaluation tasks require thinking in the third person to evaluate the actions of others, which may include inferring others’ intentions to judge the permissibility of their actions. It remains to be investigated whether such an overlap with ToM regions would occur for moral response decision tasks. Sevinc and Spreng’s (2014) recent systematic review of brain processes underlying moral cognition found activation in the default mode network. They compared brain activity for active vs. passive judgements and found that active judgements showed more activity in the temporoparietal junction (TPJ), angular gyrus, and temporal pole compared to passive viewing. Within the active domain however, they did not distinguish between moral response decision judgements and moral evaluation judgements, so it still remains to be investigated whether brain activity differs between these two types of moral decisions.

The aims of the current systematic review and meta-analysis are twofold: (a) to investigate which brain areas consistently show increased activation when making (i) moral response decisions (MRD), or (ii) making moral evaluations (ME) compared to non-moral or neutral decisions or evaluations; and (b) to compare brain activation patterns for these two types of moral judgements to

determine shared or significant differences in brain activation. A quality assessment of the included experiments was also undertaken, something which is often omitted from ALE studies.

All neuroimaging experiments of any type of moral decision-making were systematically searched and retrieved. Eligible experiments were categorised as either response decisions, or evaluations. ALE analysis was used to assess brain areas significantly more activated for both types of moral judgement, while conjunction and contrast analyses were performed to determine areas of significant difference. This allowed for the potential discrepancy in brain activation between task-type to be considered.

2. Method

2.1. Search strategy

A systematic search was conducted to identify all neuroimaging experiments of moral decision-making. Three databases, PubMed, PsycInfo and Web of Science were searched up to March 2015 using the terms “Moral” AND “Neuroimag” OR neural OR fMRI OR functional magnetic resonance OR PET OR positron emission tomography OR MEG OR magnetoencephalography OR brain”. Where the database allowed, results were limited to humans, English language and full text articles (excluding letters, editorials etc.). This search returned 3563 results (2521 after duplicates removed) which were exported to EndNote X7. A title screen was performed to remove those obviously irrelevant, followed by an abstract screen (see Fig. 1 for PRISMA flowchart). One hundred and twenty-one references remained for full text screening, which was performed based on the eligibility criteria (Table 1). The reference list of recent systematic reviews (Bzdok et al., 2012; Sevinc & Spreng, 2014) were also screened for additional references. The initial search, title and abstract screen was carried out by BG. Full text screening was carried out by BG and PL independently and decisions were compared. Where there was a disagreement, these were discussed with reference to the inclusion criteria and a joint decision was reached.

Data extraction was performed by BG for the included experiments. Details of task type and description were extracted, along with relevant foci co-ordinates for the ALE analysis. Experiments were either categorised as MRDs or MEs based on task design. Coordinates of moral vs. non-moral/neutral conditions were extracted into an Excel file and any coordinates reported in Montreal Neurological Institute (MNI) space were converted to Talairach using the icbm2tal transformation in GingerALE 2.3.4 (Laird et al., 2010; Lancaster et al., 2007). Where an experiment did fit the inclusions criteria but did not report coordinates of moral vs. non-moral/neutral conditions, the main author was contacted via email to request these data. If there was no response after three weeks, the experiment was excluded due to lack of appropriate data to extract for ALE analysis. Where an experiment included a sample of non-typically developing adults, it was only included if coordinates for the comparison group were presented separately, and only these data were extracted for analysis. For some experiments, there was more than one moral condition (e.g., moral personal and moral impersonal) and these were collapsed together to make moral vs. non-moral for extraction purposes. Where there was more than one comparison condition to a moral condition, data were extracted for moral vs. the most neutral condition (i.e., if there was a non-moral and a neutral condition, the moral vs. neutral comparison was extracted). Different experiments had different thresholds for significant clusters, but for each experiment, coordinates were extracted for moral vs. non-moral or neutral if they met the whole brain threshold set by the authors. In some papers, authors reported coordinates under the threshold because

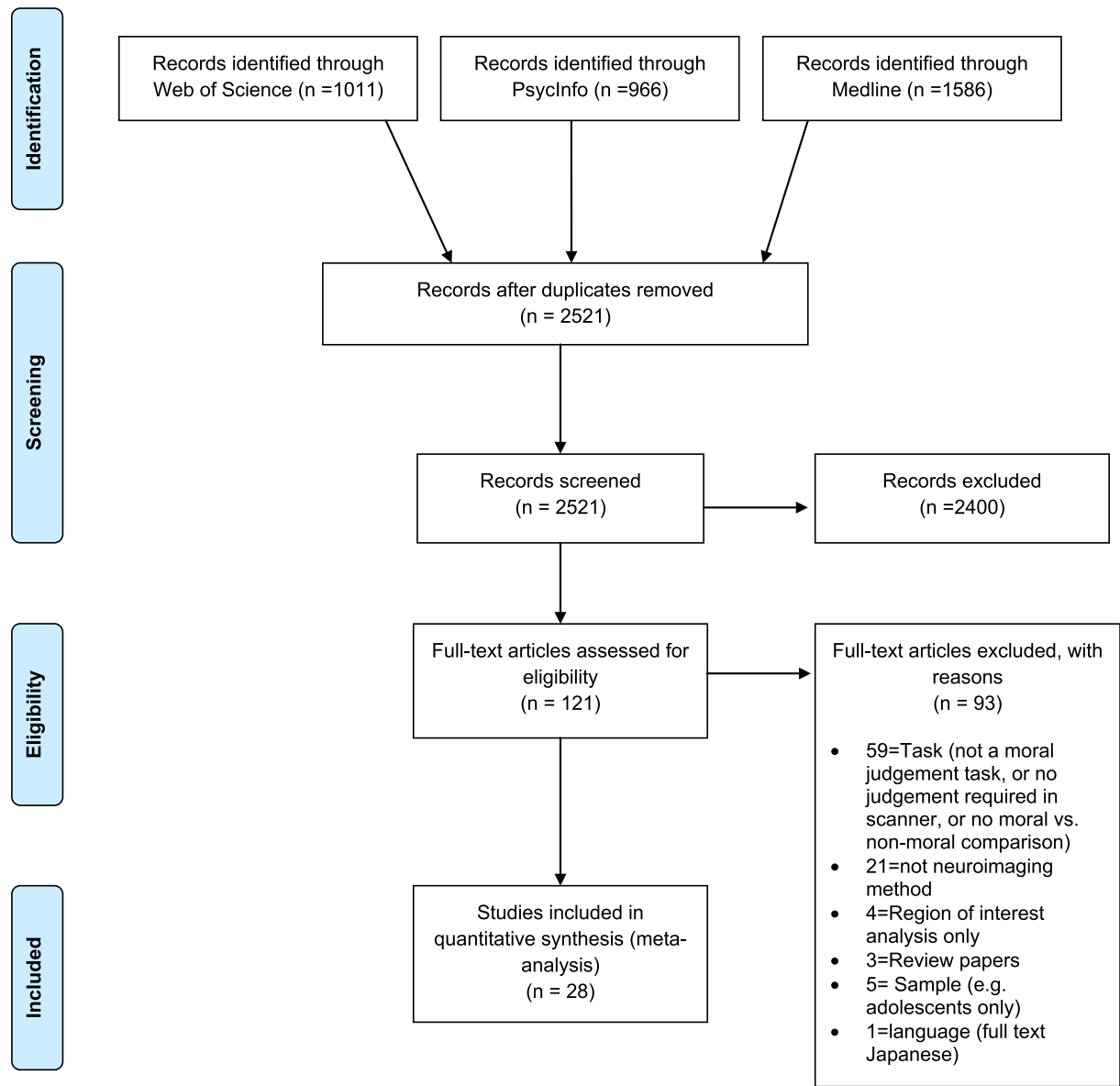


Fig. 1. PRISMA flowchart (from Moher, Liberati, Tetzlaff, & Altman, 2009).

they related to an a priori hypothesis or region of interest; these coordinates were not extracted for meta-analysis.

Table 1
Inclusion and exclusion criteria for experiments.

Inclusion criteria	Exclusion criteria
Moral decision-making task with active judgement required (response decision or evaluation)	Not moral task, or passive moral task (e.g. just viewing moral stimuli with no judgement required)
Neuroimaging method	No neuroimaging method used
Talairach or MNI coordinates reported for whole brain analysis	No coordinates reported, or region of interest analysis only
Moral task activation compared to non-moral or neutral task activation	No control task, or no moral vs. non-moral/neutral comparison
Typical/healthy adult subjects. (If experiment also includes non-typical participants, data for comparison group must be reported separately)	Coordinates reported
	Subjects with developmental or neurological disorder, psychopathy, drug dependency or children/adolescents (can include if comparison group/adult data reported separately)
English language paper	Not English language

A quality assessment tool was developed by BG and PL, based on guidelines for reporting an fMRI study (Poldrack et al., 2008), using a binary scale (1 = evidence reported, 0 = no evidence reported/unclear/not explicit; see Supplementary Material). Experiments scoring 0–10 were classed as low quality, 11–20 classed as medium quality and 21–30 classed as high quality. BG performed quality assessment for all included experiments and PL performed quality assessment on 20% of included experiments independently.

2.2. Analysis

ALE analysis is a commonly used method for coordinate based meta-analysis. This method assesses the patterns of activation foci reported in different experiments, to establish where in the brain convergence is higher than would be expected if foci were normally distributed throughout the brain (Eickhoff, Bzdok, Laird, Kurth, & Fox, 2012; Eickhoff et al., 2009; Turkeltaub et al., 2012), taking sample sizes of experiments into account. ALE analysis was performed using GingerALE 2.3.4 on the x, y, z coordinates of moral vs. non-moral or neutral conditions. Firstly, ALE analysis

was performed for all ME experiments and then for all MRD experiments. A conjunction analysis was then performed to find shared brain activation for ME and MRD judgements. Contrast analyses were performed to assess differences in brain activation (MRD-ME and ME-MRD).

Conjunction analysis comparing the results of the present ALE to results from recent ALE systematic reviews was not possible without knowledge of exactly which foci had been extracted for each included experiment of previous reviews, and was also not practical due to differences in inclusion and exclusion criteria. The results of the present ALE were instead compared visually with results from two recent reviews (Bzdok et al., 2012; Sevinc & Spreng, 2014) where appropriate, and outlined in the discussion. Results from these previous reviews were reported in MNI coordinates, so for an easier comparison with our results we transformed their reported coordinates to Talairach coordinates using the using the icbm2tal transformation in GingerAle 2.3.4 (Laird et al., 2010; Lancaster et al., 2007) and then labelled the areas using Talairach client (Lancaster et al., 1997, 2000).

3. Results

After full text screening, 28 separate experiments were eligible for inclusion, with a total of 271 foci from 642 participants. All experiments used fMRI, 10 used a MRD task and 18 used a ME task. Table 2 shows the main characteristics of the included experiments.

ALE analysis was performed for all ME experiments and all MRD experiments, *cluster-level* = 0.05, 1000 *permutations*, *p* = 0.001. Conjunction and contrast analyses were then performed, *p* = 0.01, 1000 *permutations*, *minimum cluster* = 200mm³, to assess shared and divergent brain activation between the two task types. Table 3 shows the results of ALE analysis, and Fig. 2 shows the largest significant clusters of brain activation found for each ALE analysis. All coordinates are reported in Talairach space.

Six significant clusters of activation were found across the ME experiments (18 experiments, 174 foci, 383 participants): two in the left medial frontal gyrus (MFG), the left superior temporal gyrus (STG), left cingulate gyrus (CG), right STG and right MFG. Six significant clusters were found across the MRD experiments (10 experiments, 97 foci, 259 participants): left middle temporal gyrus (MTG), left precuneus, right MFG, right MTG, right inferior frontal gyrus (IFG) and left caudate. Conjunction analysis revealed three clusters of shared activation for both moral task types: the left MTG, left cingulate gyrus and left MFG. A contrast analysis of MEs-MRDs did not find any significant clusters. However, a contrast analysis of MRDs-MEs found three significant clusters: the right MTG, right precuneus, and left MTG.

Quality assessment indicated that 20 experiments were high quality, eight were medium quality and none were low quality (see Supplementary Material). The medium quality experiments did not report as much information as the high quality experiments. Analyses included all experiments regardless of quality, but issues regarding the quality of included experiments are outlined in the discussion and should be taken into consideration when interpreting the results. Agreement between BG and PL for quality assessment was *k* = 1 (*p* = 0.14), ICC = 0.88 (*p* = 0.005).

4. Discussion

4.1. Interpretation of significant clusters of activation

This systematic review and meta-analysis builds on previous reviews in the field by differentiating between MRD judgements and ME judgements, to assess similarities and differences in

Table 2
Characteristics of included experiments.

Reference	Subjects	Conditions	No. foci	Moral stimuli	Moral task instructions/response format
<i>Moral response decisions</i> Borg, Hynes, Van Horn, Grafton, and Simnett-Armstrong (2006)	24	Moral vs. non-moral	7	Written moral scenarios	Yes/No button press to "Is it wrong to do it?" and "Would you do it?" after each scenario
Chiong et al. (2013)	16	Moral personal vs. non-moral	3	Modified Greene (2001) dilemmas presented as written text and also audible narration	Yes/No button press to "Would you do X in order to X?" after each dilemma
FeldmanHall, Sokol-Hessner, Van Bavel, and Phelps (2014)	38	Moral (difficult and easy) vs. non-moral (difficult and easy)	10	Written moral scenarios	Yes/No button press to "Do you do (proposed response)?" after each scenario
Han et al. (2014)	16	Moral (personal and impersonal) vs. neutral	28	Modified Greene et al. (2001) dilemmas, moral-personal and moral-impersonal, presented as text	Yes/no button press to "Is it appropriate for you to do (proposed solution)?"
Harrison et al. (2012)	73	Moral vs. nondilemma	11	Vignettes of modified Greene et al. (2001) dilemmas. Participants familiarised with vignettes within 1 week of scanning session using illustrations and examiner-read descriptions. Voice-prompts in scanner	Yes/no finger raise to "Would you do X?" after each dilemma
Pujol et al. (2012)	22	Moral vs. nondilemma	14	As Harrison et al. (2012)	As Harrison et al. (2012)
Schneider et al. (2012)	28	Moral vs. baseline	4	Written moral dilemmas. Social-ethical dilemmas and individual gain vs. collective losses	Agree or disagree with proposed solution using button press after each dilemma
Sommer et al. (2010)	12	Moral vs. neutral	6	Written stories describing moral conflict	"What should I do?" Choose between two response alternatives after each dilemma with a button press
Sommer et al. (2014)	16	Moral vs. social	6	As Sommer et al. (2010)	As Sommer et al. (2010)
Verdejo-Garcia et al. (2014)	14	Moral vs. nondilemma	8	As Harrison et al. (2012)	As Harrison et al. (2012)

(continued on next page)

Table 2 (continued)

Reference	Subjects	Conditions	No. foci	Moral stimuli	Moral task instructions/response format
<i>Moral evaluations</i>					
Avram et al. (2013)	16	Moral vs. esthetic	8	Written moral statements	Decide whether each moral statement could be considered 'right' by pressing button
Avram et al. (2014)	16	Moral (1st and 3rd person) vs. non-moral	10	Written moral statements	Rate each statement as 'right' or 'wrong' using button press
Bahnemann et al. (2010)	25	Moral vs. physical	10	Animated stimuli of two people in a social interaction. Protagonist's behaviour was either in accordance with or in violation of a social norm	Yes/no button press to "Is the protagonist violating a norm?" after each moral dilemma
de Achával et al. (2013)	13	Moral vs. non-moral	6	Modified version of Moral Dilemmas Test (Greene et al., 2001) presented as text	Yes/ no button press to "Would you consider it appropriate to...?" after each dilemma
Harada et al. (2009)	18	Moral vs. gender judgement	9	Written stories of protagonist performing good or bad deed	Yes/no button press to "Is the protagonist's behaviour morally bad?"
Harenski, Antonenko, Shane, and Kiehl (2008)	28	Moral vs. non-moral	15	Moral pictures from IAPS and media, depicting unpleasant social scenes indicating a moral violation	Determine if the picture represented a moral violation and rate the severity of the violation on scale of 1–5 using button press to stop rating bar
Harenski et al. (2012)	36	Moral vs. neutral	7	Moral pictures from IAPS and media, depicting unpleasant social scenes indicating a moral violation	Rate moral violation severity of picture on scale of 1–5 using button press to stop rating bar
Harenski et al. (2014)	46	Moral vs. neutral	23	Moral pictures from IAPS and media, depicting unpleasant social scenes indicating a moral violation	Rate severity of moral transgression in pictures from 1 to 5 using button press to stop rating bar
Heekeren, Wartenburger, Schmidt, Schwintowski, and Villringer (2003)	8	Moral vs. semantic	9	Written sentences (German)	Judge whether actions described in sentences were "appropriate" or "inappropriate" with button press
Heekeren et al. (2005)	12	Moral vs. semantic	8	Written sentences (German) containing bodily harm or not	Judge if actions described in moral sentences were "appropriate" or "inappropriate" with button press
Moll et al. (2001)	10	Moral vs. factual	10	Moral sentences presented aurally	Think about each statement and judge whether they are 'right' or 'wrong'. No overt judgement in scanner
Moll et al. (2002)	7	Moral vs. non-moral neutral	2	Written moral statements	Covertly judge each sentence as 'right' or 'wrong'. No overt judgement in scanner
Parkinson et al. (2011)	30	Moral (harmful, dishonest and disgust) vs. neutral	30	Written moral scenarios	Decide if main character's actions were 'wrong' or 'not wrong' with button press
Prehn et al. (2008)	23	Socio-normative vs. grammatical	6	Written sentences	Decide if action described was a social norm violation or not with yes/no button press
Reniers et al. (2012)	24	Moral vs. non-moral	6	Written scenarios of moral decision-making	Yes/no to "Is it ok for X to do this?" using button press
Robertson et al. (2007)	16	Moral (care/justice) vs. neutral nonmoral	5	Written business case scenario describing workday of a fictional marketing research analyst	Press button when identify an important point or issue
Schleim, Spranger, Erk, and Walter (2011)	40	Moral vs. personal	6	Written stories adapted from moral issues in media	Had to judge if behaviour in stories was right from a moral point of view. Decide between 'yes, rightly' and 'no, not rightly' using button press
Takahashi et al. (2008)	15	Moral (beauty and depravity) vs. neutral	4	Written sentences (Japanese)	Read sentences silently and rate according to how moral/immoral or praiseworthy/blameworthy the events were (no overt judgement in scanner)

Number of subjects is the number included in the analysis that was extracted for this review (e.g., number of subjects in control group sample) not necessarily the total number of subjects in the experiment. Harenski et al. (2014) coordinates for control group were sent by main author after email request.

Table 3

Significant clusters of activation for moral evaluations, moral response decisions, and conjunction and convergence analysis.

	#	Volume (mm ³)	X	y	z	Cerebrum	Label	Brodman area
Moral evaluations	1	3296	−6	44	20	Left	MFG	9
	2	2176	−44	−56	18	Left	STG	39
	3	2024	−2	−56	26	Left	CG	31
	4	744	50	6	−20	Right	STG	38
	5	384	2	54	2	Right	MFG	10
	6	296	−4	48	−6	Left	MFG	10
Moral response decisions	1	1968	−44	−64	20	Left	MTG	39
	2	1928	−2	−60	30	Left	Precuneus	7
	3	1512	2	44	36	Right	MFG	6
	4	1248	44	−60	24	Right	MTG	39
	5	296	36	28	−12	Right	IFG	47
	6	264	−12	4	12	Left	Caudate	–
Conjunction of moral evaluations and moral response decisions	1	712	−44	−60	18	Left	MTG	19
	2	680	−2	−56	28	Left	CG	31
	3	312	−6	44	40	Left	MFG	8
Moral response decisions-moral evaluations	1	904	42.8	−56	21	Right	MTG	39
	2	408	3	−64	30	Right	Precuneus	31
	3	264	−40	−66	16	Left	MTG	39

X, y, z coordinates are reported in Talairach space and refer to the maximum value of each cluster. Moral evaluation and moral response decision ALE analyses performed using cluster-level = 0.05, 1000 permutations, $p = 0.001$. Conjunction and contrast analysis performed using $p = 0.01$, 1000 permutations, minimum volume = 200 mm³. Labels and Brodmann areas generated by GingerALE 2.3.4. MFG = medial frontal gyrus, STG = superior temporal gyrus, CG = Cingulate gyrus, MTG = middle temporal gyrus, IFG = inferior frontal gyrus.

patterns of brain activation between these two types of moral decisions. The ALE analyses found three significant clusters of shared brain activation for both task types: the left MTG, CG and MFG. Contrast analysis revealed that MRDs additionally activated the right MTG, right precuneus and left MTG. These findings show that making one's own moral judgements about what to do in a moral dilemma is associated with increased activation of differing brain areas, as we predicted.

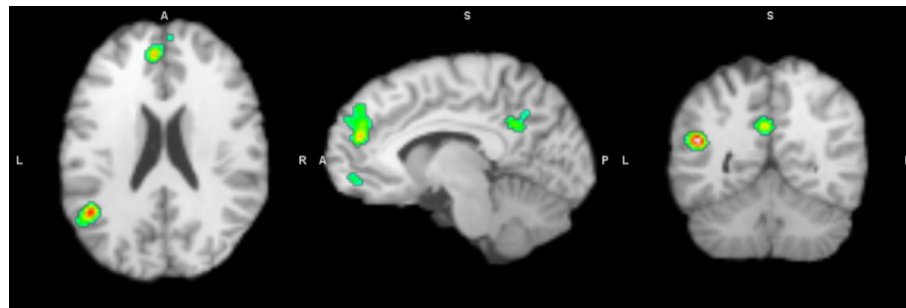
The brain region which has been most commonly implicated in moral decision-making, based on neuroimaging and lesion studies is the vmPFC. This region is not precisely defined in the literature but usually refers to any brain areas in the ventromedial frontal lobe, and BA's 10, 11, 24, 25 and 32 (Nieuwenhuis & Takashima, 2011). We only found a significant cluster of activation of this region for MEs (cluster 5 and 6, MFG BA 10) and also the adjacent BA 9 (cluster 1, MFG), rather than MRDs, and this region did not remain significant in the conjunction analysis, probably because it was the smallest of the clusters found for MEs. The lack of a significant cluster of activation in the vmPFC for MRDs highlights that most previous conclusions about brain activation for moral decision-making have been made based on ME tasks; further research on the involvement of this region for MRDs is needed, as the current review only identified 10 relevant MRD experiments.

The three clusters of significant shared activation for both moral task type (ME and MRD) - the left MTG, left CG, and left MFG - are also involved in other processes, so are not unique or specific to making moral judgements. Such a view that there is no 'moral brain' suggests that many processes such as attention, working memory, emotion recognition, empathic arousal and retrieval of relevant schemas may be involved when making moral judgements, thus many brain areas related to various domains are likely to be recruited. All three significant clusters were found in the left hemisphere, which is involved in language (Springer et al., 1999) so this may reflect the fact that most of the tasks involve language processing. It has been found that perceptual decisions engaged the left hemisphere of the MFG (Talati & Hirsch, 2005) and that the MTG is involved in multimodal semantic processing (Visser, Jefferies, Embleton, & Ralph, 2012) with the left MTG being the core component of the semantic network (Wei et al., 2012). The cluster of activation in the cingulate gyrus was found in BA 31,

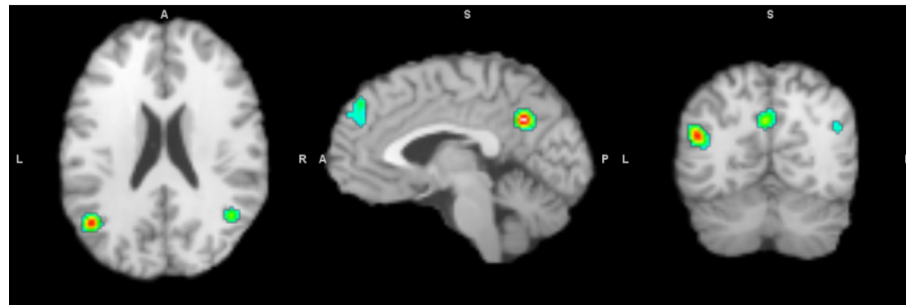
which is part of the posterior cingulate cortex and has been found to show an increase in activation when judging the valence of emotional words (Maddock, Garrett, & Buonocore, 2003); increased activation of this area for both types of moral task may reflect processing of written emotional stimuli.

Relative to ME tasks, MRDs were found to additionally activate the left and right MTG and the right precuneus. These findings support our hypothesis that MRDs will show increased activation of more self-referential brain areas. The precuneus, a brain region more highly developed in humans than other animals, is involved in higher order cognitive processes including self-processing and consciousness (Cavanna & Trimble, 2006) and egocentric spatial processing (Fretton et al., 2014). The MTG also showed an increase in activation for MRDs but not MEs, suggesting it may play a role in making one's own decisions, particularly the right MTG which was not a significant cluster in the conjunction analysis. While activation of the right precuneus in MRD tasks may reflect increased self-referential processing compared to when making MEs of other's behaviour, it may just reflect differences between the moral task types. The right precuneus is associated with metaphor comprehension (Mashal, Vishne, & Laor, 2014) and verbal creative thinking (Chen et al., 2015), so activation of this region during MRD tasks may reflect the fact that these tasks tend to involve dilemmas that are not real life (e.g., choosing to kill one or five people), thus may require more abstract thinking about unfamiliar situations.

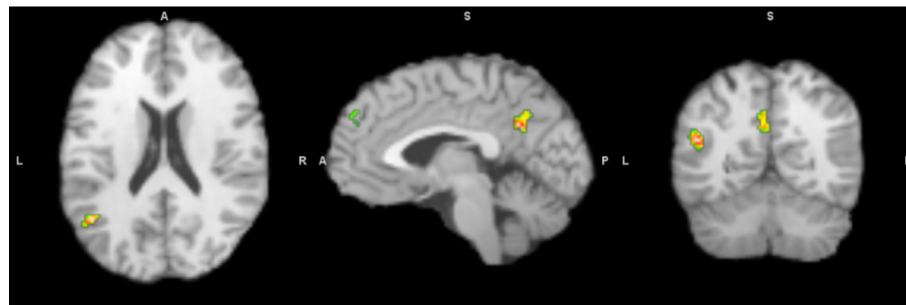
We hypothesised that the ME tasks would show increased activation of more ToM related areas than MRD tasks, as they involve thinking about the mental states of others to judge moral behaviour. This hypothesis was not supported, as the contrast analysis for MEs-MRDs did not reveal any significant clusters. The right temporoparietal junction (rTPJ) has been suggested as an area important for ToM (Saxe & Powell, 2006). The rTPJ is a vaguely defined area but is also referred to as Brodmann Area (BA) 39 (Bzdok et al., 2013). Contrary to our hypothesis, the ALE analysis revealed that the rTPJ (BA 39, MTG) showed significantly increased activation across the MRD tasks (cluster 4) but not across the ME tasks. The surprising finding of significant activation of this area for MRDs but not MEs suggests that ToM processes are even more involved when making one's own moral decisions than when making evaluations of others. One explanation may be that when



Moral evaluation clusters



Moral response decision clusters



Conjunction analysis: Shared activation for moral evaluation and moral response decisions



Contrast analysis: Moral response decisions-moral evaluations

Fig. 2. Brain activation maps showing significant clusters of activation. Images created in GingerAle 2.3.4, overlaid onto Colin $2 \times 2 \times 2$ template in Mango (Lancaster & Martinez 2006–2015). Brain images are axial, sagittal and coronal view of main clusters of activation for each ALE analysis. Image labels: L = left, R = right, P = posterior, A = anterior, S = superior.

thinking about what to do in a moral dilemma, individuals think about the consequences of their possible actions for others, e.g. “would my actions upset/harm someone?” ToM abilities develop with age (Wellman, Cross, & Watson, 2001) and the perspectives of others are taken into account more as egocentric bias decreases (Gibbs, 2013), and this may reflect the increased activation of the rTPJ for MRDs amongst adults found in this meta-analysis. Harenski, Harenski, Shane, and Kiehl’s (2012) ME study which we

included in our meta-analysis also included an adolescent sample, and they found that involvement of the rTPJ while viewing moral pictures increased with age. Developmental fMRI studies of MRDs are needed, to establish whether the involvement of ToM regions when making one’s own moral decisions increases with age. The hypothetical dilemmas used in the included MRD experiments involved other people, so participants may typically infer the mental states or possible mental states of others when deciding their

response. Some of the included ME tasks used did not reference other people, such as judging sentences as 'right' or 'wrong' so would not have led to participants inferring mental states of others. Contrary to our finding, Bzdok et al. (2012) found significant activation in the rTPJ for their moral cognition domain. However, experiments were only included in their moral cognition analysis if they involved participants making "appropriateness judgements on actions of one individual towards others" (p. 785) so always involved other people. The lack of rTPJ involvement for MEs in this meta-analysis may reflect the type of evaluation tasks used. Our finding suggested that not all types of moral decision-making involve ToM processes – it depends on whether the dilemma or stimuli involves other people, which can lead individuals to infer the mental states of others when considering the possible consequences of their decisions. Real life moral dilemmas are likely to involve other people, so ToM processes are likely to be involved in such decisions, and involvement may change with age.

We compared our findings to those of two recent systematic reviews of moral decision-making, Bzdok et al. (2012) and Sevinc and Spreng (2014). As previously stated, the criteria for Bzdok et al.'s (2012) moral cognition domain was that participants were required to make appropriateness judgements, which is what we have termed moral evaluations. We, therefore, compared our ALE analysis results for MEs to Bzdok et al.'s (2012) results for the moral cognition domain. Our results are fairly comparable to Bzdok et al.'s (2012) with both finding activation in the left and right MFG (BA 10, labelled by Bzdok et al. as vmPFC) and the right STG (BA 38, labelled by Bzdok et al. as the right temporal pole). In line with Bzdok et al.'s (2012) analysis, we also found a cluster of activation in the left STG (labelled by Bzdok et al. as left TPJ), though in our analysis this was BA 39 and from their analysis it was BA 22, although these are adjacent areas. As previously stated, Bzdok et al. (2012) found a cluster of activation in the rTPJ (BA 39), which we did not find. Another discrepancy between our ME activation clusters and Bzdok et al.'s (2012) for moral cognition are that they found a cluster of activation in the left amygdala, which we did not find. Also, Bzdok et al. (2012) reported activation in the precuneus, which was not found to be a cluster of significant activation for the ME experiments in our analysis, although our results did reveal activation in the adjacent cingulate gyrus. Differences between our findings and Bzdok et al.'s (2012) may be partly due to discrepancies between Talairach labels and the SPM anatomy toolbox, and also due to the differences of tasks for the included experiments; Bzdok et al. (2012) included only experiments where participants were required to make appropriateness judgements on the actions of one individual towards others while we included any ME judgement, including tasks where participants had to judge moral sentences. Our ALE analysis results for MRDs are not comparable with Bzdok et al.'s (2012) moral cognition results, supporting the finding that judging the appropriateness of others actions increases activation of different brain regions compared to when making one's own moral response decisions.

Sevinc and Spreng's (2014) systematic review compared brain activation for active and passive moral tasks. As our review only included tasks that required an active decision, we compared our findings for MRDs and MEs to Sevinc and Spreng's (2014) findings for active tasks. Our ALE results for MEs are fairly comparable to Sevinc and Spreng's (2014) ALE results for active tasks, with both finding clusters of activation in the left MFG (BA 9, labelled by Sevinc & Spreng as medial PFC and anterior superior frontal sulcus), right MFG (BA 10, labelled by Sevinc & Spreng as vmPFC and superior temporal sulcus), right STG (BA 38, labelled by Sevinc & Spreng as MTG) and the left cingulate gyrus (BA 31, labelled by Sevinc & Spreng as posterior cingulate cortex). Comparing our results for MRDs to Sevinc and Spreng's (2014) active results, there is no direct overlap but there are some similarities. Sevinc and

Spreng (2014) found a cluster of activation in the left MTG, BA 22, whereas our ALE analysis for MRDs revealed a cluster in the left MTG BA 39, adjacent to BA 22. Sevinc and Spreng (2014) found activation in the left IFG whereas we found activation in the right IFG for MRDs. Again, differences may be due to discrepancies between Talairach and MNI labelling and also differences between tasks of the included experiments. For Sevinc and Spreng's (2014) active domain, four of the included experiments were also included in our MRD domain, but we included an additional six MRD experiments, and the majority of the active experiments in Sevinc and Spreng (2014) were MEs.

4.2. Quality assessment and critique of tasks used in included experiments

As far as we are aware, this review is the first ALE meta-analysis to report on the quality of included experiments. In the absence of a pre-existing standardised quality assessment tool for an ALE systematic review, the quality assessment tool was adapted from guidelines for reporting fMRI studies (Poldrack et al., 2008). Future reviews could use this checklist, or similar quality assessment tools, and should exclude low quality experiments from ALE analysis. While the majority of included experiments in this systematic review were found to be of high quality, based on the adapted checklist, there were issues with some of the tasks used.

Firstly, there was a wide range of moral tasks used across the included experiments, and it is important to acknowledge that differences in brain activation may reflect differences in the tasks used across studies, in terms of modality and content of the moral stimuli, and the nature of the control task. The presentation modality varied, and while most tasks used written stimuli presented on screens for participants to read, some experiments, such as Bahnemann, Dziobek, Prehn, Wolf, and Heekeren (2010) used animated stimuli, with participants being asked to judge if the protagonist in the animations is violating a norm. Also, there were differences in the context and amount of detail of moral stimuli across moral tasks. Bahnemann et al.'s (2010) animations featured a social violation where a protagonist is punching the other person in the face. While this is a real life social and moral (harm) violation, it could be argued that this scenario would be more emotive if the victim was murdered. Differences in emotional engagement, paying more attention to out of the ordinary stimuli, or having to think more about scenarios not previously encountered (such as many of the life and death choices presented in moral tasks) are likely to contribute to differences across experiments.

One limitation of some of the included experiments is the nature of the task that was used as a comparison to the moral task. For some experiments, participants were asked to respond similarly across control and experimental conditions, while the stimuli varied. For example, in Parkinson et al. (2011), participants judged whether a character's actions were right or wrong within neutral or moral tasks. Differences in brain activity may therefore have been partially accounted for by differences in reading and processing moral, compared to neutral scenarios, rather than for making a moral compared to a neutral evaluation judgement, which the authors did acknowledge (p. 3166). Similar issues exist across several other experiments (Harenski et al., 2012; Harenski, Edwards, Harenski, & Kiehl, 2014; Han, Glover, & Jeong, 2014; Moll, Eslinger, & de Oliveira-Souza, 2001; Moll et al., 2002; Reniers et al., 2012). Several of the included experiments used the same moral task, classified as a MRD task, where participants were asked "would you do X?" after reading a scenario about Mr. Jones' dilemma (Harrison et al., 2012; Pujol et al., 2008; Verdejo-Garcia et al., 2014). For the control condition, participants were asked to recall the correct answer to the non-dilemma vignette, which they had been familiarised with before the scanner task, e.g., "will he go

to the beach?” The control task used in these experiments was, therefore, a recall rather than a decision task, thus the results show brain activation differences for recall vs. a MRD judgement, rather than brain activation for making a moral as opposed to a non-moral decision. For some of the included experiments, the control task may not have been an appropriate comparison for moral tasks (i.e., not a non-moral or neutral judgement task), so we should be cautious about the significant moral decision-making peak coordinates from these experiments.

4.3. Limitations

There are several limitations to this review and meta-analysis. Firstly, only 10 MRD experiments were identified from the literature, and 15 is the minimum recommended number of experiments for ALE contrast analysis (Laird et al., 2010; Lancaster et al., 2007). While significant clusters were still found for MRD and ME judgements, and the findings are novel, the results should be interpreted with caution as the MRD clusters are only based on 10 experiments. Secondly, there were some experiments where it was ambiguous as to whether the task was a MRD or ME task. We categorised the experiments based on the authors' claims, and the type of question participants responded to (evaluation or response decision), but difficulty categorising some of the experiments highlights the lack of consistency amongst moral tasks used in fMRI experiments. Thirdly, some of the tasks appeared to lack ecological validity as they did not seem to reflect how moral decisions are made in real life situations. We recommend that future neuroimaging experiments use more real life scenarios for assessing moral decision-making, for example, everyday scenarios that people are more likely to encounter than life or death situations. Finally, to ensure comparability across studies, adolescents were excluded. The conclusions drawn, therefore, only apply to adults. Further neuroimaging studies focusing on children and adolescents would help answer questions about moral development and the developmental pattern of the neural correlates of moral decisions.

5. Conclusion

This is the first systematic review of moral decision-making to explicitly acknowledge the different types of moral decisions, and to compare brain activity for the two main types, MRDs and MEs. Findings from the ALE analysis show that making one's own moral judgements about what to do in a moral dilemma involves increased activation of additional brain areas compared to judging the moral actions of others, suggesting different processes may be involved. Making one's own decisions appears to involve an extended brain network, incorporating self-referential regions which do not show an increase in activation when making moral evaluations of others. Most previous conclusions about moral decision-making have been based on moral evaluation tasks; further neuroimaging studies employing moral response decisions tasks of real life scenarios are needed before we can be confident about which brain areas are needed for making one's own, every day moral response decisions.

Conflict of interest

All authors declare they have no conflict of interest.

Acknowledgements

Peter E. Langdon is supported by a National Institute for Health Research Postdoctoral Fellowship (Grant Reference: NIHR-PDF-2011-04-040). This article presents independent research funded

by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the National Health Service, the National Institute for Health Research or the Department of Health. The National Institute for Health Research, the National Health Service, or the Department of Health had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandc.2016.07.007>.

References

- Avram, M., Gutyrchik, E., Bao, Y., Pöppel, E., Reiser, M., & Blautzik, J. (2013). Neurofunctional correlates of esthetic and moral judgments. *Neuroscience Letters*, 534, 128–132.
- Avram, M., Hennig-Fast, K., Bao, Y., Pöppel, E., Reiser, M., Blautzik, J., et al. (2014). Neural correlates of moral judgments in first- and third-person perspectives: Implications for neuroethics and beyond. *BMC Neuroscience*, 15(1), 39.
- Bahnmann, M., Dziobek, I., Prehn, K., Wolf, I., & Heekeren, H. R. (2010). Sociotopy in the temporoparietal cortex: Common versus distinct processes. *Social Cognitive and Affective Neuroscience*, 5(1), 48–58.
- Blair, J., Marsh, A., Finger, E., Blair, K., & Luo, J. (2006). Neuro-cognitive systems involved in morality. *Philosophical Explorations*, 9(1), 13–27.
- Borg, J. S., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: An fMRI investigation. *Journal of Cognitive Neuroscience*, 18(5), 803–817.
- Bzdok, D., Langner, R., Schilbach, L., Jakobs, O., Roski, C., Caspers, S., et al. (2013). Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. *Neuroimage*, 81, 381–392.
- Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A., Langner, R., et al. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, 217(4), 783–796.
- Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*, 129(3), 564–583.
- Chen, Q.-L., Xu, T., Yang, W.-J., Li, Y.-D., Sun, J.-Z., Wang, K.-C., et al. (2015). Individual differences in verbal creative thinking are reflected in the precuneus. *Neuropsychologia*, 75, 441–449.
- Chiong, W., Wilson, S. M., d'Esposito, M., Kayser, A. S., Grossman, S. N., Poorzand, P., et al. (2013). The salience network causally influences default mode network activity during moral reasoning. *Brain: A Journal of Neurology*, 136(6), 1929–1941.
- Christensen, J., Flexas, A., Calabrese, M., Gut, N. K., & Gomila, A. (2014). Moral judgment reloaded: A moral dilemma validation study. *Frontiers in Psychology*, 5.
- Christensen, J., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience & Biobehavioral Reviews*, 36(4), 1249–1264.
- de Achaval, D., Villarreal, M. F., Salles, A., Bertomeu, M. J., Costanzo, E. Y., Goldschmidt, M., et al. (2013). Activation of brain areas concerned with social cognition during moral decisions is abnormal in schizophrenia patients and unaffected siblings. *Journal of Psychiatric Research*, 47(6), 774–782.
- Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *Neuroimage*, 59(3), 2349–2361.
- Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, 30(9), 2907–2926.
- FeldmanHall, O., Sokol-Hessner, P., Van Bavel, J. J., & Phelps, E. A. (2014). Fairness violations elicit greater punishment on behalf of another than for oneself. *Nature Communications*, 5.
- Freton, M., Lemogne, C., Bergouignan, L., Delaveau, P., Lehericy, S., & Fossati, P. (2014). The eye of the self: Precuneus volume and visual perspective during autobiographical memory retrieval. *Brain Structure and Function*, 219(3), 959–968.
- Fumagalli, M., & Priori, A. (2012). Functional and clinical neuroanatomy of morality. *Brain*, 135(7), 2006–2021.
- Gibbs, J. C. (2013). *Moral development and reality: Beyond the theories of Kohlberg, Hoffman, and Haidt*. Oxford University Press.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105–2108.
- Han, H., Glover, G. H., & Jeong, C. (2014). Cultural influences on the neural correlate of moral decision making processes. *Behavioural Brain Research*, 259, 215–228.

- Harada, T., Itakura, S., Xu, F., Lee, K., Nakashita, S., Saito, D. N., et al. (2009). Neural correlates of the judgment of lying: A functional magnetic resonance imaging study. *Neuroscience Research*, 63(1), 24–34.
- Harenski, C. L., Antonenko, O., Shane, M. S., & Kiehl, K. A. (2008). Gender differences in neural mechanisms underlying moral sensitivity. *Social Cognitive and Affective Neuroscience*, 3(4), 313–321.
- Harenski, C. L., Edwards, B. G., Harenski, K. A., & Kiehl, K. A. (2014). Neural correlates of moral and non-moral emotion in female psychopathy. *Frontiers in Human Neuroscience*, 8.
- Harenski, C. L., Harenski, K. A., Shane, M. S., & Kiehl, K. A. (2012). Neural development of mentalizing in moral judgment from adolescence to adulthood. *Developmental Cognitive Neuroscience*, 2(1), 162–173.
- Harrison, B. J., Pujol, J., Soriano-Mas, C., Hernández-Ribas, R., López-Solà, M., Ortiz, H., et al. (2012). Neural correlates of moral sensitivity in obsessive-compulsive disorder. *JAMA Psychiatry*, 69(7), 741–749.
- Heekeren, H. R., Wartenburger, I., Schmidt, H., Prehn, K., Schwintowski, H.-P., & Villringer, A. (2005). Influence of bodily harm on neural correlates of semantic and moral decision-making. *Neuroimage*, 24(3), 887–897.
- Heekeren, H. R., Wartenburger, I., Schmidt, H., Schwintowski, H.-P., & Villringer, A. (2003). An fMRI study of simple ethical decision-making. *NeuroReport: For Rapid Communication of Neuroscience Research*, 14(9), 1215–1219.
- Laird, A. R., Robinson, J. L., McMillan, K. M., Tordesillas-Gutiérrez, D., Moran, S. T., Gonzales, S. M., et al. (2010). Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: Validation of the Lancaster transform. *Neuroimage*, 51(2), 677–683.
- Lancaster, J., & Martinez, M. (Copyright 2006–2015). *Mango*. Research Imaging Institute: University of Texas Health Science Center.
- Lancaster, J., Rainey, L., Summerlin, J., Freitas, C., Fox, P., Evans, A., et al. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping*, 5(4), 238.
- Lancaster, J., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., et al. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping*, 28(11), 1194–1205.
- Lancaster, J., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10(3), 120–131.
- Maddock, R. J., Garrett, A. S., & Buonocore, M. H. (2003). Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Human Brain Mapping*, 18(1), 30–41.
- Marazziti, D., Baroni, S., Landi, P., Ceresoli, D., & Dell'Osso, L. (2013). The neurobiology of moral sense: Facts or hypotheses? *Annals of General Psychiatry*, 12(1), 6.
- Mashal, N., Vishne, T., & Laor, N. (2014). The role of the precuneus in metaphor comprehension: evidence from an fMRI study in people with schizophrenia and healthy participants. *Frontiers in Human Neuroscience*, 8, 818.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–269.
- Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiuolo, P. A., et al. (2002). The neural correlates of moral sensitivity: A functional magnetic resonance imaging investigation of basic and moral emotions. *The Journal of Neuroscience*, 22(7), 2730–2736.
- Moll, J., Eslinger, P. J., & de Oliveira-Souza, R. (2001). Frontopolar and anterior temporal cortex activation in a moral judgment task. *Arquivos de Neuro-Psiquiatria*, 59(3-B), 657–664.
- Nieuwenhuis, I. L. C., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behavioural Brain Research*, 218(2), 325–334.
- Parkinson, C., Sinnott-Armstrong, W., Koralus, P. E., Mendelovici, A., McGeer, V., & Wheatley, T. (2011). Is morality unified? Evidence that distinct neural systems underlie moral judgments of harm, dishonesty, and disgust. *Journal of Cognitive Neuroscience*, 23(10), 3162–3180.
- Poldrack, R. A., Fletcher, P. C., Henson, R. N., Worsley, K. J., Brett, M., & Nichols, T. E. (2008). Guidelines for reporting an fMRI study. *Neuroimage*, 40(2), 409–414.
- Prehn, K., Wartenburger, I., Mériaux, K., Scheibe, C., Goodenough, O. R., Villringer, A., et al. (2008). Individual differences in moral judgment competence influence neural correlates of socio-normative judgments. *Social Cognitive and Affective Neuroscience*, 3(1), 33–46.
- Pujol, J., Reixach, J., Harrison, B. J., Timoneda-Gallart, C., Vilanova, J. C., & Pérez-Alvarez, F. (2008). Posterior cingulate activation during moral dilemma in adolescents. *Human brain mapping*, 29(8), 910–921.
- Pujol, J., Batalla, I., Contreras-Rodríguez, O., Harrison, B. J., Pera, V., Hernández-Ribas, R., et al. (2012). Breakdown in the brain network subserving moral judgment in criminal psychopathy. *Social Cognitive and Affective Neuroscience*, 7(8), 917–923.
- Raine, A., & Yang, Y. (2006). Neural foundations to moral reasoning and antisocial behavior. *Social Cognitive and Affective Neuroscience*, 1(3), 203–213.
- Reniers, R. L. E. P., Corcoran, R., Völlm, B. A., Mashru, A., Howard, R., & Liddle, P. F. (2012). Moral decision-making, ToM, empathy and the default mode network. *Biological Psychology*, 90(3), 202–210.
- Robertson, D., Snarey, J., Ousley, O., Harenski, K., Bowman, F. D., Gilkey, R., et al. (2007). The neural processing of moral sensitivity to issues of justice and care. *Neuropsychologia*, 45(4), 755–766.
- Saxe, R., & Powell, L. J. (2006). It's the thought that counts specific brain regions for one component of theory of mind. *Psychological Science*, 17(8), 692–699.
- Schleim, S., Spranger, T. M., Erk, S., & Walter, H. (2011). From moral to legal judgment: The influence of normative context in lawyers and other academics. *Social Cognitive and Affective Neuroscience*, 6(1), 48–57.
- Schneider, K., Pauly, K. D., Gossen, A., Mevissen, L., Michel, T. M., Gur, R. C., et al. (2012). Neural correlates of moral reasoning in autism spectrum disorder. *Social cognitive and affective neuroscience*, nss051.
- Sevinc, G., & Spreng, R. N. (2014). Contextual and perceptual brain processes underlying moral cognition: a quantitative meta-analysis of moral reasoning and moral emotions.
- Smetana, J. G. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 119–155). Mahwah, New Jersey London: Lawrence Erlbaum Associates.
- Sommer, M., Meinhardt, J., Rothmayr, C., Döhl, K., Hajak, G., Rupprecht, R., et al. (2014). Me or you? Neural correlates of moral reasoning in everyday conflict situations in adolescents and adults. *Social Neuroscience*, 9(5), 452–470.
- Sommer, M., Rothmayr, C., Döhl, K., Meinhardt, J., Schwerdtner, J., Sodian, B., et al. (2010). How should I decide? The neural correlates of everyday moral reasoning. *Neuropsychologia*, 48(7), 2018–2026.
- Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S., et al. (1999). Language dominance in neurologically normal and epilepsy subjects. *Brain*, 122(11), 2033–2046.
- Takahashi, H., Kato, M., Matsuura, M., Koeda, M., Yahata, N., Suhara, T., et al. (2008). Neural correlates of human virtue judgment. *Cerebral Cortex*, 18(8), 1886–1891.
- Talati, A., & Hirsch, J. (2005). Functional specialization within the medial frontal gyrus for perceptual go/no-go decisions based on “what”, “when”, and “where” related information: An fMRI study. *Journal of cognitive neuroscience*, 17(7), 981–993.
- Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between judgment and choice of action in moral dilemmas. *Frontiers in Psychology*, 4.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.
- Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human Brain Mapping*, 33(1), 1–13.
- Verdejo-García, A., Contreras-Rodríguez, O., Fonseca, F., Cuenca, A., Soriano-Mas, C., Rodríguez, J., et al. (2014). Functional alteration in frontolimbic systems relevant to moral judgment in cocaine-dependent subjects. *Addiction Biology*, 19(2), 272–281.
- Visser, M., Jefferies, E., Embleton, K. V., & Ralph, M. A. L. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: Distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience*, 24(8), 1766–1778.
- Wei, T., Liang, X., He, Y., Zang, Y., Han, Z., Caramazza, A., et al. (2012). Predicting conceptual processing capacity from spontaneous neuronal activity of the left middle temporal gyrus. *The Journal of Neuroscience*, 32(2), 481–489.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655–684.
- Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, 7(1), 1–10.